Unconventional kinetochore kinases KKT2 and KKT3 have a unique zinc finger that promotes their kinetochore localization

Gabriele Marcianò†, Olga O. Nerusheva†, Midori Ishii†, and Bungo Akiyoshi*

Department of Biochemistry, University of Oxford, Oxford, OX1 3QU, UK

†These authors contributed equally to this work

*For correspondence: bungo.akiyoshi@bioch.ox.ac.uk

Abstract

Chromosome segregation in eukaryotes is driven by the kinetochore, a macromolecular protein complex that assembles onto centromeric DNA and binds spindle microtubules. Cells must tightly control the number and position of kinetochores so that all chromosomes assemble a single kinetochore. A central player in this process is the centromere-specific histone H3 variant CENP-A, which localizes specifically within centromeres and promotes kinetochore assembly. However, CENP-A is absent from several eukaryotic lineages including kinetoplastids, a group of evolutionarily divergent eukaryotes that have an unconventional set of kinetochore proteins. It remains unknown how kinetoplastids specify kinetochore positions or promote kinetochore assembly in the absence of CENP-A. Here we studied two homologous kinetoplastid kinases (KKT2 and KKT3) that localize constitutively at centromeres. KKT2 and KKT3 central domains were sufficient for centromere localization in *Trypanosoma brucei*. Crystal structures of the KKT2 central domain from two divergent kinetoplastids revealed a unique zinc finger domain, which promotes its kinetochore localization in *T. brucei*. Mutations in the equivalent zinc finger domain of KKT3 abolished its kinetochore localization and function. This study lays the foundation for understanding the mechanism of kinetochore specification and assembly in kinetoplastids.

Introduction

The kinetochore is the macromolecular protein complex that drives chromosome segregation during mitosis and meiosis in eukaryotes. Its fundamental functions are to bind DNA and spindle microtubules (Musacchio and Desai, 2017). In most eukaryotes, kinetochores assemble within a single chromosomal region, called the centromere. While components of spindle microtubules are highly conserved across eukaryotes (Wickstead and Gull, 2011; Findeisen et al., 2014), centromere DNA is known to evolve rapidly (Henikoff et al., 2001). It is critical that a single kinetochore is formed per chromosome and its position is maintained between successive cell divisions. A key player involved in this kinetochore specification process is the centromere-specific histone H3 variant CENP-A, which is found in most sequenced eukaryotic genomes (Talbert et al., 2009). CENP-A localizes specifically at centromeres throughout the cell cycle and recruits HJURP, a specific chaperone that incorporates CENP-A onto centromeres (Black and Cleveland, 2011; McKinley and Cheeseman, 2016; Stankovic and Jansen, 2017). Besides CENP-A, components of the constitutive centromere-associated network (CCAN) also localize at centromeres constitutively. CENP-A-containing nucleosomes are recognized by CCAN components, which in turn recruit the KNL1-Mis12-Ndc80 (KMN) network that has microtubule-binding activities. In addition to these structural kinetochore proteins, several protein kinases are known to localize at mitotic kinetochores, including Cdk1, Aurora B, Bub1, Mps1, and Plk1 (Cheeseman and Desai, 2008). These protein kinases regulate various aspects of mitosis, including kinetochore assembly, error correction, and the spindle checkpoint (Carmena et al., 2012; London and Biggins, 2014; Hara and Fukagawa, 2018).

Kinetoplastids are evolutionarily divergent eukaryotes that are defined by the presence of a unique organelle called the kinetoplast that contains a cluster of mitochondrial DNA (d'Avila-Levy et al., 2015). Centromere positions have been mapped in three kinetoplastids, *Trypanosoma brucei* (Obado et al., 2007), *Trypanosoma cruzi* (Obado et al., 2005), and *Leishmania* (Garcia-Silva et al., 2017), which all have regional centromeres. There is no specific DNA sequence that is common to all centromeres in each organism, suggesting that kinetoplastids also determine their kinetochore positions in a sequence-independent manner. However, none of CENP-A or any other canonical structural kinetochore proteins has been identified in kinetoplastids (Lowell and Cross, 2004;

Berriman et al., 2005; Aslett et al., 2010). They instead have unique kinetochore proteins, such as KKT1–25 (Akiyoshi and Gull, 2014; Nerusheva and Akiyoshi, 2016; Nerusheva et al., 2019) and KKIP1–12 (D'Archivio and Wickstead, 2017; Brusini et al., 2019) in *T. brucei*. Among these proteins, six proteins localize at centromeres throughout the cell cycle (KKT2, KKT3, KKT4, KKT20, KKT22, and KKT23), implying close association with centromeric DNA. KKT2 and KKT3 are homologous proteins and have three domains conserved among kinetoplastids: a protein kinase domain that is classified as unique among known eukaryotic kinase subfamilies (Parsons et al., 2005), a central domain of unknown function, and divergent polo boxes (DPB) (Figure 1 and 2). Presence of an N-terminal kinase domain and a C-terminal DPB suggests that KKT2 and KKT3 likely share common ancestry with polo-like kinases (Nerusheva and Akiyoshi, 2016). Interestingly, a protein kinase domain is not present in any constitutively-localized kinetochore protein in other eukaryotes, making these protein kinases a unique feature of kinetoplastid kinetochores. In addition to the three domains that are highly conserved among kinetoplastids, AT-hook and SPKK DNA-binding motifs are found in some species, suggesting that these proteins are located close to DNA (Akiyoshi and Gull, 2014). Although RNAi-mediated knockdown of KKT2 or KKT3 leads to growth defects (Akiyoshi and Gull, 2014; Jones et al., 2014), little is known about their molecular function. In this report, we have revealed a unique zinc finger in the KKT2 and KKT3 central domain, which is important for their kinetochore localization and function in *T. brucei*.

Results

The central domain of KKT3 is able to localize at centromeres in *T. brucei*

To understand how KKT3 localizes at centromeres, we determined which domain was responsible for its centromere localization by expressing in *T. brucei* procyclic form (insect stage) cells, a series of truncated versions of *Tb*KKT3, fused with a GFP-tagged nuclear localization signal peptide (GFP-NLS) (Figure 1A). We found that the N-terminal protein kinase domain or the C-terminal DPB of *Tb*KKT3 did not localize at centromeres. In contrast, *Tb*KKT3$^{594–1058}$ and *Tb*KKT3$^{594–811}$ that contain the central region formed kinetochore-like dots (Figure 1B). *Tb*KKT3$^{594–728}$ also formed dots, suggesting that AT-hook and SPKK DNA-binding motifs (residues 771–780) are dispensable for

centromere localization in this assay. These results revealed the importance of the central domain for the recruitment of *Tb*KKT3 onto centromeres.

Multiple domains of KKT2 are able to localize at centromeres in *T. brucei*

We next performed a similar analysis for *Tb*KKT2 and found that its central domain (residues 562–677) was able to localize at centromeres (Figure 2A, B). Therefore, the central domains of both *Tb*KKT2 and *Tb*KKT3 are sufficient for centromere localization in this assay. In a previous study, we showed that ectopically-expressed *Tb*KKT2 DPB (1024–1260) localized at centromeres (Nerusheva and Akiyoshi, 2016). The present study confirmed this result and also identified another region (residues 672–1030) that localized at centromeres (Figure 2A, B). These results show that *Tb*KKT2 has multiple domains that are sufficient for centromere localization in this assay. Based on our previous finding that *Tb*KKT2 co-immunoprecipitated with a number of kinetochore proteins (Akiyoshi and Gull, 2014), we reasoned that these fragments might localize at centromeres by interacting with other kinetochore proteins. To test this possibility, we immunoprecipitated *Tb*KKT2 fragments and performed mass spectrometry to identify co-purifying proteins. We found that *Tb*KKT2 DPB co-purified with several kinetochore proteins (KKT1, KKT6, KKT7, and KKT8), which was abolished in the *Tb*KKT2 W1048A mutant that did not localize at kinetochores (Figure 2C and Table S1) (Nerusheva and Akiyoshi, 2016). Similarly, *Tb*KKT2$^{672–1030}$ co-purified with several kinetochore proteins, with KKIP1 being the top hit (Figure 2D and Table S1). These results suggest that, when ectopically expressed in trypanosomes, *Tb*KKT2$^{672–1030}$ and *Tb*KKT2 DPB likely localize at centromeres by interacting with other kinetochore proteins. Because those co-purifying kinetochore proteins (e.g. KKT1, KKT6, KKT7, KKT8, KKIP1) localize at kinetochores only transiently, it is likely that the constitutively-localized *Tb*KKT2 protein recruits them using *Tb*KKT2$^{672–1030}$ and DPB domains. Furtheremore, the fact that KKIP1 was detected in the immunoprecipitates of *Tb*KKT2, but not other kinetochore proteins (Nerusheva et al., 2019), raises a possibility that KKIP1 is recruited to kinetochores by directly interacting with the 672–1030 region of *Tb*KKT2. By contrast, the central domains of *Tb*KKT2 and *Tb*KKT3 did not robustly co-purify with other kinetochore proteins (Table S1).

The *Bodo saltans* KKT2 central domain adopts a unique zinc finger structure

To gain insights into how the central domains of KKT2 and KKT3 localize specifically at centromeres, we expressed and purified recombinant *Tb*KKT2 and *Tb*KKT3 proteins for their structure determination by X-ray crystallography. Our attempts to purify the *Tb*KKT3 central domain were unsuccessful, but we managed to express and purify from *E. coli* the central domain of KKT2 from several kinetoplastids, including *Bodo saltans* (a free-living kinetoplastid (Jackson et al., 2016)) and *Perkinsela* (endosymbiotic kinetoplastids (Tanifuji et al., 2017)) (Figure S1). We successfully obtained crystals of the *Bs*KKT2 central domain (residues 572–668), which allowed us to determine its structure to 1.8 Å resolution by zinc single-wavelength anomalous dispersion (Zn-SAD) phasing (Figure 3 and Table 1). Our analysis revealed the presence of two distinct zinc fingers: an N-terminal zinc finger (referred to as Znf1) consisting of 2 β-sheets (where β-strands 1, 4 and 5 comprise the first β-sheet, and β-strands 2 and 3 comprise the second β-sheet) and one α-helix, and a C-terminal zinc finger (referred to as Znf2) consisting of one β-sheet (comprising β-strands 6 and 7) and one α-helix (Figure 3A, B). *Bs*KKT2 Znf1 coordinates two zinc ions and Znf2 coordinates one zinc ion.

A structural homology search using the DALI server (Holm and Laakso, 2016) indicated that *Bs*KKT2 Znf1 has weak structural similarity to proteins that have C1 domains (Table S2). C1 domains were originally discovered as lipid-binding modules in protein kinase Cs (PKCs) and are characterized by the $HX_{12}CX_2CXnCX_2CX_4HX_2CX_7C$ motif (Colón-González and Kazanietz, 2006; Das and Rahman, 2014). C1 domains are classified into a typical C1 domain, which binds diacylglycerol or phorbol esters, and an atypical C1 domain, which is not known to bind ligands. The closest structural homolog of *Bs*KKT2 Znf1 was the atypical C1 domain of the Vav1 protein (RMSD 2.7 Å across 52 Cα). Although *Bs*KKT2 Znf1 and the Vav1 C1 domain share some structural similarity, their superposition revealed fundamental differences (Figure 4). Coordination of one zinc ion in *Bs*KKT2 Znf1 occurs via the N-terminal residues Cys 580 and His 584, while that in the Vav1 C1 domain occurs via the N-terminal His 516 and C-terminal Cys 564. More importantly, KKT2 Znf1 does not have the $HX_{12}CX_2CXnCX_2CX_4HX_2CX_7C$ motif that is present in all C1 domains. Therefore,

KKT2 Znf1 and C1 domains are distinct domains, and their structural similarity is likely to be a product of convergent evolution.

Structural analysis of *Bs*KKT2 Znf2 revealed a classical C2H2 zinc finger domain (Table S3). C2H2 zinc fingers are known to bind DNA, RNA, or protein (Krishna et al., 2003; Brayer and Segal, 2008). In most known cases, two or more C2H2 zinc fingers are used to recognize specific DNA sequences (Wolfe et al., 2000). This recognition is typically achieved by specific interactions between the side chain of residues in positions -1, 2, 3, and 6 in the recognition α-helix (where -1 is the residue immediately preceding the α-helix) and DNA bases (Wolfe et al., 2000). Notably, some proteins with a single zinc finger can recognize specific DNA sequences (Omichinski et al., 1997; Dathan et al., 2002). *Bs*KKT2 Znf2 structure consists of one C2H2 domain (-1: Ser 653, 2: Thr 655, 3: Lys 656, 6: Tyr 659). The sequence alignment of *Bs*KKT2 Znf2 shows that residues at the position -1 and 3 are highly conserved in trypanosomatids, while those at position 2 and 6 are not (Figure 6A).

Znf1 is conserved in *Perkinsela* KKT2a

We next asked whether the central domain structure is conserved among kinetoplastids. *Perkinsela* is a highly divergent endosymbiotic kinetoplastid that lives inside *Paramoeba* (Tanifuji et al., 2017). Homology search identified three proteins that have similarity to KKT2 and KKT3. Interestingly, similarities among these *Perkinsela* proteins are higher than those between them and KKT2 or KKT3 in other kinetoplastids (Figure S2). Because these *Perkinsela* proteins overall have higher sequence similarity to KKT2 than KKT3, we call them *Pk*KKT2a (XU18_4017), *Pk*KKT2b (XU18_0308), and *Pk*KKT2c (XU18_4564). *Pk*KKT2c does not have a kinase domain, like KKT20 in other kinetoplastids (Nerusheva and Akiyoshi, 2016). Our sequence alignment suggests that *Pk*KKT2a and *Pk*KKT2b have a Znf1-like domain but do not have a Znf2-like C2H2 domain (Figure S2).

We determined the crystal structure of the *Pk*KKT2a central domain (residues 551–679) at 2.9 Å resolution by Zn-SAD phasing (Figure S1 and Table 2), which confirmed the presence of a Znf1-like structure: 2 β-sheets (residues 551–647), followed by an extended C-terminal α-helix (residues 648–679) (Figure 5A, B). The Znf1-like domain of *Pk*KKT2a overlaps closely with that of

*Bs*KKT2 (RMSD: 0.79 Å across 39 Cα), with the exception of some differences being localized to the loop insertion and the absence of a C2H2 domain in the *Pk*KKT2a central domain (Figure 5C), consistent with our sequence analysis (Figure S2). Taken together, our structures have revealed that the KKT2 central domain contains a zinc finger domain that is conserved between *Bodo saltans* and *Perkinsela*. Given the sequence similarity of KKT2 between *Bodo saltans* and other trypanosomatids (Figure 6A), it is likely that the unique zinc finger structure is conserved among kinetoplaitids.

Zinc finger Znf1 of KKT2 is important for long-term viability in *T. brucei*

To examine the functional relevance of the KKT2 Znf1 and Znf2 zinc fingers, we tested their mutants in *T. brucei*. Because KKT2 has multiple domains that can independently localize at kinetochores (Figure 2), we first expressed mutants in our ectopic expression of the central domain (*Tb*KKT2$^{562–677}$). We found that mutations in *Tb*KKT2 Znf1 Zn-coordinating residues (C576A, H580A, C597A, C600A, C616A, C619A) all abolished kinetochore localization (Figure 6A, B). In contrast, similar mutations in *Tb*KKT2 Znf2 (C640A C643A) did not affect the localization.

To gain insights into how KKT2 Znf1 may promote kinetochore localization, we analysed conservation and electrostatic potential of *Bs*KKT2 Znf1 surface residues, to identify possible patches that may be involved in this process. Our analysis revealed a highly conserved acidic patch, centered around residue *Bs*KKT2 D626 (Figure 3C, D, and Figure S3). Interestingly, this aspartic acid is strictly conserved in all KKT2 and KKT3 proteins (Figure S2). To test the importance of this residue, we mutated the corresponding residue and found that *Tb*KKT2 D622A and *Tb*KKT2 D622E mutants failed to localize at kinetochores (Figure 6B). Taken together, our results show that Znf1, but not Znf2, is important for the kinetochore localization of the *Tb*KKT2 central domain.

To test the importance of the *Tb*KKT2 central domain for cell viability, we performed rescue experiments. We replaced one allele of *Tb*KKT2 with an N-terminally YFP-tagged *Tb*KKT2 construct that has either wild-type or mutant versions of the central domain, and performed RNAi against the 5'UTR of the *Tb*KKT2 transcript to knock down the untagged allele of *Tb*KKT2 (Figure S4) (Ishii and Akiyoshi, 2019). As expected, mutants in Znf1 (C576A and D622A) and Znf2 (C640A C643A) both localized at kinetochores (Figure 6C). Upon induction of RNAi, however, the Znf1 mutants

failed to support normal growth after day 4, while the Znf2 mutants rescued the growth defects. These data confirm the importance of Znf1 for the function of *Tb*KKT2 *in vivo*.

Localization of KKT3 depends on the central domain in *T. brucei*

While *Tb*KKT2 has multiple domains that can promote centromere localization, *Tb*KKT3 has only the central domain that was able to localize at kinetochores (Figure 1). Consistent with this result, mutating *Tb*KKT3 residues that align with Zn-coordinating histidine or cysteine residues in the Znf1 of KKT2 abolished the kinetochore localization of the ectopically-expressed full length *Tb*KKT3 protein (Figure 7A, B). We also found that the conserved aspartic acid *Tb*KKT3 D692 was essential for kinetochore localization. In contrast, mutations in *Tb*KKT3 Znf2 (C707A C710A) did not affect the kinetochore localization.

We next performed rescue experiments by replacing one allele of *Tb*KKT3 with a C-terminally YFP-tagged construct that has either wild-type or mutant versions of the central domain, and performed RNAi against the 3'UTR of *Tb*KKT3 to knock down the untagged allele of *Tb*KKT3 (Figure S4). We first confirmed that *Tb*KKT3 Znf1 mutants (C668A C671A and D692A) were unable to localize at kinetochores, while the *Tb*KKT3 Znf2 mutant (C707A C710A) localized normally. Upon induction of RNAi, Znf1 mutants failed to rescue the growth defect, showing that kinetochore localization is essential for the *Tb*KKT3 function (Figure 7C). In contrast, the *Tb*KKT3 Znf2 mutant supported normal cell growth. These data show that the *Tb*KKT3 Znf1 zinc finger is essential for the localization and function of *Tb*KKT3, similarly to what was observed for *Tb*KKT2 Znf1.

Discussion

A major open question concerning the unconventional kinetochores of kinetoplastids is how these organisms specify kinetochore positions without CENP-A. In other words, how do kinetochore proteins localize specifically at centromeres in a sequence-independent manner? Towards addressing this important question, we examined two constitutively localized kinetochore kinases KKT2 and KKT3, which have sequence similarity with polo-like kinases (Nerusheva and Akiyoshi, 2016). In addition to an N-terminal protein kinase domain and a C-terminal polo box domain present in polo-

like kinases, KKT2 and KKT3 have a central domain that is highly conserved among kinetoplastids. By expressing fragments of KKT2 and KKT3 in trypanosomes, we first established that their central domains can localize specifically at centromeres. The crystal structure of the *Bodo saltans* KKT2 central domain then revealed a unique structure, which consists of two distinct zinc fingers (Znf1 and Znf2). It is likely that the central domain of *T. brucei* KKT2 has a similar structure based on high sequence similarity between *Bs*KKT2 and *Tb*KKT2 proteins. Importantly, mutational analyses of KKT2 in *T. brucei* revealed that Znf1 was important for the localization of the central domain, while Znf2 was not. Furthermore, although full-length *Tb*KKT2 Znf1 mutants localized at kinetochores (likely due to interactions with other kinetochore proteins via other domains), they were not fully functional. Taken together, these data have established that Znf1 is essential for the function of *Tb*KKT2, which is supported by the presence of Znf1, but not Znf2, in *Perkinsela* KKT2a. We speculate that Znf2 has affinity for DNA and stabilizes the localization of KKT2 at centromeres in *T. brucei* and *Bodo saltans*. In agreement with this hypothesis, putative DNA-binding SPKK motifs (Suzuki, 1989) are present right after Znf2 in many kinetoplastids (Figure 6A), while *Perkinsela* KKT2b has an AT-hook motif within the Znf1 domain (Figure S2). It remains unclear whether the structure of the central domain is conserved between KKT2 and KKT3. Nonetheless, our functional studies of *Tb*KKT3 mutants showed that the Znf1 equivalent domain was also essential for the kinetochore localization and function, while the Znf2 equivalent domain was not, showing that the functional importance of Znf1 is conserved in KKT3. It will be important to obtain KKT3 central domain structures to reveal structural similarity or difference between KKT2 and KKT3.

The fact that KKT2/3 central domains manage to localize specifically at centromeres suggests that they are able to recognize something special at centromeres. What might be a unique feature at centromeres in kinetoplastids that lack CENP-A? Histone variants are one possibility. *T. brucei* has four histone variants, H2AZ, H2BV, H3V, and H4V. However, none of them is specifically enriched at centromeres (Lowell and Cross, 2004; Lowell et al., 2005; Siegel et al., 2009), and histone chaperones did not co-purify with any kinetochore protein (Akiyoshi and Gull, 2014). Alternatively, there might exist certain post-translational modifications on histones or DNA specifically at centromeres (e.g. phosphorylation, methylation, or acetylation). KKT2 Znf1 has a highly conserved

acidic patch, which might act as a “reader” for such modifications. Although there is no known histone or DNA modification that occurs specifically at centromeres, KKT2/3 have a protein kinase domain and KKT23 has a Gcn5-related N-acetyltransferase (GNAT) domain (Nerusheva et al., 2019). It will be important to examine whether these enzymatic domains are important for recruiting the central domain and identify their substrates. Another unique feature at centromeres is the presence of kinetochore proteins, which could potentially recruit newly synthesized kinetochore components by direct protein-protein interactions. Finally, it is important to note that it remains unclear whether kinetoplastid kinetochores build upon nucleosomes. It is formally possible that the KKT2/3 central domains directly bind DNA and form a unique environment at centromeres. Understanding how the KKT2/3 central domains localize specifically at centromeres will be key to elucidating the mechanism of how kinetoplastids specify kinetochore positions in the absence of CENP-A.

Materials and Methods

Tryp cells and plasmids, microscopy, immunoprecipitation, and mass spectrometry

All trypanosome cell lines, plasmids, primers, and synthetic DNA used in this study are listed in Table S4. All trypanosome cell lines used in this study were derived from *T. brucei* SmOxP927 procyclic form cells (TREU 927/4 expressing T7 RNA polymerase and the tetracycline repressor to allow inducible expression) (Poon et al., 2012). Cells were grown at 28 ˚C in SDM-79 medium supplemented with 10% (v/v) heat-inactivated fetal calf serum (Brun and Schönenberger, 1979). Endogenous YFP tagging was performed using the pEnT5-Y vector (Kelly et al., 2007). Inducible expression of GFP-NLS fusion proteins or hairpin RNAi was performed using pBA310 (Nerusheva and Akiyoshi, 2016).

To make pBA1711 (KKT3 3'UTR hairpin RNAi), BAG95 was digested with HindIII/BamHI and subcloned into the pBA310 HindIII/BamHI sites. To make pBA1807, 4–3174bp of KKT3 coding sequence and 250bp of 3'UTR were amplified with primers BA2351/BA2352 and BA2353/BA2354, digested with HindIII/NotI and NotI/SpeI, respectively, and were cloned into the pEnT5-Y using HindIII/SpeI sites. Site-directed mutagenesis was performed using primers and template plasmids listed in Table S4. All constructs were sequence verified.

Plasmids linearized by NotI were transfected into trypanosomes by electroporation into an endogenous locus (pEnT5-Y derivatives) or 177bp repeats on minichromosomes (pBA310 derivatives). To obtain endogenously-tagged strains (pEnT5-Y derivatives), transfected cells were selected by the addition of 25 µg/ml hygromycin and cloned by dispensing dilutions into 96-well plates. Clones that express mutant versions of KKT2 or KKT3 were screened by Sanger sequencing of genomic DNA. pBA310 derivatives were selected by the addition of 5 µg/ml phleomycin. Expression of GFP-NLS fusion proteins (pBA310 derivatives) was induced by the addition of doxycycline (10 ng/ml). RNAi was induced by the addition of doxycycline (1 µg/ml).

Fluorescence microscopy, immunoprecipitation of GFP-fusion proteins, and mass spectrometry were performed essentially as described previously (Nerusheva and Akiyoshi, 2016; Ishii and Akiyoshi, 2019). Proteins identified with at least two peptides were considered as significant and shown in Table S1.

Multiple sequence alignment

Protein sequences and accession numbers for KKT2 and KKT3 homologs were retrieved from TriTryp database (Aslett et al., 2010), Wellcome Sanger Institute (https://www.sanger.ac.uk/), or UniProt (UniProt Consortium, 2019). Search for KKT2/3 homologs in *Perkinsela* was done using hmmsearch on its predicted proteome using manually prepared KKT2/3 hmm profiles (HMMER version 3.0 (Eddy, 1998)). Multiple sequence alignment was performed with MAFFT (L-INS-i method, version 7) (Katoh and Standley, 2013) and visualized with the Clustalx coloring scheme in Jalview (version 2.10) (Waterhouse et al., 2009).

Protein expression and purification

Multiple sequence alignment together with secondary structure predictions of the KKT2 central domain were used to design constructs in *Bodo saltans* and *Perkinsela*. To make pBA1660 (*Bs*KKT2$^{572–668}$ with an N-terminal TEV-cleavable hexahistidine (His_6) tag), the central domain of *Bodo saltans* KKT2 (accession number BSAL_50690) was amplified from BAG50 (a synthetic DNA that encodes *Bodo saltans* KKT2, codon optimized for expression in Sf9 insect cells (Table S4)) with primers BA2117/BA2118 and cloned into the RSFDuet-1 vector (Novagen) using BamHI/EcoRI sites with an NEBuilder HiFi DNA Assembly Cloning Kit (NEB) according to the manufacturer's instructions. To make pBA1139 (His_6-*Pk*KKT2$^{551–679}$), the central domain of *Perkinsela* CCAP 1560/4 KKT2a (accession number XU18_4017) was amplified from BAG48 (a synthetic DNA that encodes *Perkinsela* KKT2a, codon optimized for expression in Sf9 insect cells (Table S4)) with primers BA1569/BA1570 and cloned into RSFDuet-1 using BamHI/EcoRI sites with an In-Fusion HD Cloning Plus kit (Takara Bio). Recombinant proteins were expressed in BL21(DE3) *E. coli* cells at 20 ˚C using auto induction media (Formedium) (Studier, 2005).

Briefly, 500 mL of cells were grown at 37 ˚C in 2.5 L flasks at 300 rpm until OD_{600} of 0.2–0.3 and then cooled down to 20 ˚C overnight (2 L for *Bs*KKT2 and 6 L for *Pk*KKT2a). Cells were harvested by centrifugation and resuspended in 50 ml per litre of culture of lysis buffer (25 mM Hepes pH 7.5, 150 mM NaCl, 1 mM TCEP, 10 mM imidazole, and 1.2 mM PMSF). Proteins were

extracted by mechanical cell disruption using a French press (1 passage at 20,000 PSI) and the resulting lysate was centrifuged at 48,384 g for 30 min at 4 ˚C. Clarified lysate was incubated with 5 mL TALON beads (Takara Bio), washed with 150 mL lysis buffer and eluted in 22 mL of elution buffer (25 mM Hepes pH 7.5, 150 mM NaCl, 1 mM TCEP, and 250 mM imidazole) in a gravity column, followed by TEV treatment for the removal of the His_6 tag. Salt concentration of the sample was subsequently reduced to 50 mM NaCl using buffer A (50 mM Hepes pH 7.5, and 1 mM TCEP) and the sample was loaded onto a 5 mL HiTrap Heparin HP affinity column (GE healthcare) pre-equilibrated with 5% buffer B (50 mM Hepes pH 7.5, 1 M NaCl, and 1 mM TCEP) on an ÄKTA pure 25 system. Protein was eluted by using a gradient from 0.05 to 1 M NaCl, and protein-containing fractions were combined, concentrated with an Amicon stirred cell using an ultrafiltration disc with 10 kDa cut-off (Merck), and then loaded onto a HiPrep Superdex 75 16/60 size exclusion chromatography column (GE healthcare) pre-equilibrated with 25 mM Hepes pH 7.5, 150 mM NaCl, and 1 mM TCEP. Fractions containing the KKT2 central domain were pooled together, concentrated with an Amicon stirred cell using an ultrafiltration disc with 10 kDa cut-off, and stored at -80 ˚C. Protein concentration was measured by Bradford assay.

Crystallization

Both *Bs*KKT2 and *Pk*KKT2a central domain crystals were optimized at 4 ˚C in sitting drop vapour diffusion experiments in 48-well plates, using drops of overall volume 400 nL, mixing protein and mother liquor in a 3:1 protein:mother liquor ratio. *Bs*KKT2 central domain crystals grew from the protein at 26 mg/mL and mother liquor 40% PEG 400, 0.2 mM $(NH_4)_2SO_4$, and 100 mM Tris-HCl pH 8. The 40% PEG400 in the mother liquor served as the cryoprotectant when flash-cooling the crystals by plunging into liquid nitrogen. *Pk*KKT2a central domain crystals grew from the protein at 13 mg/mL and mother liquor 19% MPD, 50 mM Hepes pH 7.5, and 10 mM $MgCl_2$. The crystals were briefly transferred into a cryoprotecting solution of 30% MPD, 50 mM Hepes pH 7.5, 10 mM $MgCl_2$ prior to flash-cooling.

Data collection and structure determination

X-ray diffraction data from a *Bs*KKT2 central domain crystal were collected at the I04 beamline at the Diamond Light Source (Harwell, UK) at the Zinc K-edge wavelength (λ=1.28297 Å). A set of 1441 images were processed in space group I222 using the Xia2 pipeline (Winter, 2010), with DIALS for indexing and integration (Winter et al., 2018) and AIMLESS for scaling (Evans and Murshudov, 2013) to 1.8 Å resolution. Initial 3 Zn atoms were localized by interpreting the anomalous difference Patterson, single anomalous dispersion (SAD) phases were estimated using Crank2 (Skubák and Pannu, 2013), and an initial model was built with BUCCANEER (Cowtan, 2006). The structure was completed by several cycles of alternating model building in Coot (Emsley et al., 2010) and refinement in autoBUSTER (Blanc et al., 2004; Bricogne et al., 2017).

*Pk*KKT2a central domain X-ray diffraction data were collected at the I03 beamline at Diamond Light Source (Harwell, UK) also at the zinc K-edge (λ=1.28272 Å) and processed using the autoPROC pipeline (Vonrhein et al., 2011) using XDS (Kabsch, 2010) for indexing/integration and AIMLESS (Evans and Murshudov, 2013) for scaling, to a resolution of 3.8 Å. Two initial Zn positions were determined by interpreting the anomalous difference Patterson, and single anomalous dispersion (SAD) phases were estimated using Crank2 (Skubák and Pannu, 2013) and SHARP (Vonrhein et al., 2007) in space group $P6_4$. An initial model was manually built in Coot and refined once with RosettaMR (Terwilliger et al., 2012). The structure was completed by several cycles of alternating model building in Coot (Emsley et al., 2010) and refinement in autoBUSTER (Blanc et al., 2004; Bricogne et al., 2017).

A higher resolution dataset was collected from a *Pk*KKT2a central domain crystal at the I24 beamline at Diamond Light Source (Harwell, UK), at a wavelength of λ=0.9686 Å. Data were processed using Xia2 pipeline (Winter, 2010), DIALS (Winter et al., 2018), and AIMLESS (Evans and Murshudov, 2013) in space group $P6_4$ to a resolution of 2.9 Å. The model obtained from the 3.8 Å dataset was used for further model building and refinement with autoBuster (Bricogne et al., 2017) and Coot (Emsley et al., 2010).

All the images were made with Pymol (Schrödinger LLC, Portland, OR) and CCP4mg (McNicholas et al., 2011). Topology diagrams were generated using TopDraw (Bond, 2003). Protein

coordinates have been deposited in the RCSB Protein Data Bank (http://www.rcsb.org/) with accession codes 6TLY (*Bodo saltans* KKT2) and 6TLX (*Perkinsela* KKT2a).

Acknowledgments

We thank Pietro Roversi and Matt Higgins for advice, crystallography facility manager Edward Lowe, as well as Pietro Roversi, Matt Higgins, Danny Huang, and Patryk Ludzia for comments on the manuscript. Midori Ishii was supported by a long-term fellowship from the TOYOBO Biotechnology Foundation. Bungo Akiyoshi was supported by a Wellcome Trust Senior Research Fellowship (grant no. 210622/Z/18/Z) and the European Molecular Biology Organization Young Investigator Program. The authors declare that no competing interests exist.

Author contributions

Gabriele Marcianò purified recombinant proteins and solved crystal structures. Olga Nerusheva expressed KKT2 and KKT3 truncations and mutants in trypanosomes and performed immunoprecipitation and mass spectrometry. Midori Ishii performed rescue experiments. Bungo Akiyoshi expressed KKT2 and KKT3 truncations and mutants in trypanosomes. Gabriele Marcianò, Midori Ishii, and Bungo Akiyoshi wrote the manuscript.

Author ORCIDs

Gabriele Marcianò, https://orcid.org/0000-0003-2720-652X

Midori Ishii, https://orcid.org/0000-0001-9597-9458

Bungo Akiyoshi, https://orcid.org/0000-0001-6010-394X

Data availability

All data generated during this study are included in the manuscript and supplementary data. Protein coordinates have been deposited in the RCSB Protein Data Bank (http://www.rcsb.org/) with accession codes 6TLY (*Bodo saltans* KKT2) and 6TLX (*Perkinsela* KKT2a).

References

Akiyoshi, B., and Gull, K. (2014). Discovery of unconventional kinetochores in kinetoplastids. Cell *156*, 1247–1258.

Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., and Ben-Tal, N. (2016). ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res. *44*, W344-350.

Aslett, M., Aurrecoechea, C., Berriman, M., Brestelli, J., Brunk, B. P., Carrington, M., Depledge, D. P., Fischer, S., Gajria, B., Gao, X., et al. (2010). TriTrypDB: a functional genomic resource for the Trypanosomatidae. Nucleic Acids Res. *38*, D457-462.

Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D. C., Lennard, N. J., Caler, E., Hamlin, N. E., Haas, B., et al. (2005). The genome of the African trypanosome *Trypanosoma brucei*. Science *309*, 416–422.

Black, B. E., and Cleveland, D. W. (2011). Epigenetic centromere propagation and the nature of CENP-A nucleosomes. Cell *144*, 471–479.

Blanc, E., Roversi, P., Vonrhein, C., Flensburg, C., Lea, S. M., and Bricogne, G. (2004). Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. Acta Crystallogr. D Biol. Crystallogr. *60*, 2210–2221.

Bond, C. S. (2003). TopDraw: a sketchpad for protein structure topology cartoons. Bioinformatics *19*, 311–312.

Brayer, K. J., and Segal, D. J. (2008). Keep your fingers off my DNA: protein-protein interactions mediated by C2H2 zinc finger domains. Cell Biochem. Biophys. *50*, 111–131.

Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O. S., Vonrhein, C., et al. (2017). BUSTER version 2.10.3. Cambridge, United Kingdom: Global Phasing Ltd.

Brun, R., and Schönenberger (1979). Cultivation and in vitro cloning or procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. Short communication. Acta Trop. *36*, 289–292.

Brusini, L., D'Archivio, S., McDonald, J., and Wickstead, B. (2019). Ndc80/Nuf2-like protein KKIP1 connects a stable kinetoplastid outer kinetochore complex to the inner kinetochore and responds to metaphase tension. bioRxiv, 764829.

Carmena, M., Wheelock, M., Funabiki, H., and Earnshaw, W. C. (2012). The chromosomal passenger complex (CPC): from easy rider to the godfather of mitosis. Nat. Rev. Mol. Cell Biol. *13*, 789–803.

Cheeseman, I. M., and Desai, A. (2008). Molecular architecture of the kinetochore-microtubule interface. Nat. Rev. Mol. Cell Biol *9*, 33–46.

Colón-González, F., and Kazanietz, M. G. (2006). C1 domains exposed: from diacylglycerol binding to protein-protein interactions. Biochim. Biophys. Acta *1761*, 827–837.

Cowtan, K. (2006). The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr. D Biol. Crystallogr. *62*, 1002–1011.

D'Archivio, S., and Wickstead, B. (2017). Trypanosome outer kinetochore proteins suggest conservation of chromosome segregation machinery across eukaryotes. J. Cell Biol. *216*, 379–391.

Das, J., and Rahman, G. M. (2014). C1 domains: structure and ligand-binding properties. Chem. Rev. *114*, 12108–12131.

Dathan, N., Zaccaro, L., Esposito, S., Isernia, C., Omichinski, J. G., Riccio, A., Pedone, C., Di Blasio, B., Fattorusso, R., and Pedone, P. V. (2002). The *Arabidopsis* SUPERMAN protein is able to specifically bind DNA through its single Cys 2 –His 2 zinc finger motif. Nucleic Acids Res *30*, 4945–4951.

Eddy, S. R. (1998). Profile hidden Markov models. Bioinformatics *14*, 755–763.

Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

Evans, P. R., and Murshudov, G. N. (2013). How good are my data and what is the resolution? Acta Cryst D *69*, 1204–1214.

Findeisen, P., Mühlhausen, S., Dempewolf, S., Hertzog, J., Zietlow, A., Carlomagno, T., and Kollmar, M. (2014). Six subgroups and extensive recent duplications characterize the evolution of the eukaryotic tubulin protein family. Genome Biol Evol *6*, 2274–2288.

Garcia-Silva, M.-R., Sollelis, L., MacPherson, C. R., Stanojcic, S., Kuk, N., Crobu, L., Bringaud, F., Bastien, P., Pagès, M., Scherf, A., et al. (2017). Identification of the centromeres of *Leishmania major*: revealing the hidden pieces. EMBO Reports *18*, 1968–1977.

Hara, M., and Fukagawa, T. (2018). Kinetochore assembly and disassembly during mitotic entry and exit. Curr. Opin. Cell Biol. *52*, 73–81.

Henikoff, S., Ahmad, K., and Malik, H. S. (2001). The centromere paradox: stable inheritance with rapidly evolving DNA. Science *293*, 1098–1102.

Holm, L., and Laakso, L. M. (2016). Dali server update. Nucleic Acids Res. *44*, W351-355.

Ishii, M., and Akiyoshi, B. (2019). Unconventional kinetochore kinases KKT10/19 promote the metaphase to anaphase transition in *Trypanosoma brucei*. bioRxiv, 806224.

Jackson, A. P., Otto, T. D., Aslett, M., Armstrong, S. D., Bringaud, F., Schlacht, A., Hartley, C., Sanders, M., Wastling, J. M., Dacks, J. B., et al. (2016). Kinetoplastid Phylogenomics Reveals the Evolutionary Innovations Associated with the Origins of Parasitism. Curr. Biol. *26*, 161–172.

Jones, N. G., Thomas, E. B., Brown, E., Dickens, N. J., Hammarton, T. C., and Mottram, J. C. (2014). Regulators of *Trypanosoma brucei* cell cycle progression and differentiation identified using a kinome-wide RNAi screen. PLoS Pathog. *10*, e1003886.

Jurrus, E., Engel, D., Star, K., Monson, K., Brandi, J., Felberg, L. E., Brookes, D. H., Wilson, L., Chen, J., Liles, K., et al. (2018). Improvements to the APBS biomolecular solvation software suite. Protein Science *27*, 112–128.

Kabsch, W. (2010). XDS. Acta Crystallogr. D Biol. Crystallogr. *66*, 125–132.

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780.

Kelly, S., Reed, J., Kramer, S., Ellis, L., Webb, H., Sunter, J., Salje, J., Marinsek, N., Gull, K., Wickstead, B., et al. (2007). Functional genomics in *Trypanosoma brucei*: a collection of vectors for the expression of tagged proteins from endogenous and ectopic gene loci. Mol. Biochem. Parasitol. *154*, 103–109.

Krishna, S. S., Majumdar, I., and Grishin, N. V. (2003). Structural classification of zinc fingersSURVEY AND SUMMARY. Nucleic Acids Res *31*, 532–550.

Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., and Ben-Tal, N. (2005). ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. Nucleic Acids Res. *33*, W299-302.

d'Avila-Levy, C. M., Boucinha, C., Kostygov, A., Santos, H. L. C., Morelli, K. A., Grybchuk-Ieremenko, A., Duval, L., Votýpka, J., Yurchenko, V., Grellier, P., et al. (2015). Exploring the environmental diversity of kinetoplastid flagellates in the high-throughput DNA sequencing era. Mem. Inst. Oswaldo Cruz *110*, 956–965.

London, N., and Biggins, S. (2014). Signalling dynamics in the spindle checkpoint response. Nat. Rev. Mol. Cell Biol. *15*, 736–747.

Lowell, J. E., and Cross, G. A. M. (2004). A variant histone H3 is enriched at telomeres in *Trypanosoma brucei*. J. Cell. Sci. *117*, 5937–5947.

Lowell, J. E., Kaiser, F., Janzen, C. J., and Cross, G. A. M. (2005). Histone H2AZ dimerizes with a novel variant H2B and is enriched at repetitive DNA in *Trypanosoma brucei*. J. Cell. Sci. *118*, 5721–5730.

McKinley, K. L., and Cheeseman, I. M. (2016). The molecular basis for centromere identity and function. Nat. Rev. Mol. Cell Biol. *17*, 16–29.

McNicholas, S., Potterton, E., Wilson, K. S., and Noble, M. E. M. (2011). Presenting your structures: the CCP4mg molecular-graphics software. Acta Cryst D *67*, 386–394.

Musacchio, A., and Desai, A. (2017). A Molecular View of Kinetochore Assembly and Function. Biology (Basel) *6*, 5.

Nerusheva, O. O., and Akiyoshi, B. (2016). Divergent polo box domains underpin the unique kinetoplastid kinetochore. Open Biol *6*, 150206.

Nerusheva, O. O., Ludzia, P., and Akiyoshi, B. (2019). Identification of four unconventional kinetoplastid kinetochore proteins KKT22-25 in *Trypanosoma brucei*. Open Biol *9*, 190236.

Obado, S. O., Bot, C., Nilsson, D., Andersson, B., and Kelly, J. M. (2007). Repetitive DNA is associated with centromeric domains in *Trypanosoma brucei* but not *Trypanosoma cruzi*. Genome Biol. *8*, R37.

Obado, S. O., Taylor, M. C., Wilkinson, S. R., Bromley, E. V., and Kelly, J. M. (2005). Functional mapping of a trypanosome centromere by chromosome fragmentation identifies a 16-kb GC-rich transcriptional "strand-switch" domain as a major feature. Genome Res. *15*, 36–43.

Omichinski, J. G., Pedone, P. V., Felsenfeld, G., Gronenborn, A. M., and Clore, G. M. (1997). The solution structure of a specific GAGA factor–DNA complex reveals a modular binding mode. Nature Structural Biology *4*, 122–132.

Parsons, M., Worthey, E. A., Ward, P. N., and Mottram, J. C. (2005). Comparative analysis of the kinomes of three pathogenic trypanosomatids: *Leishmania majo*r, *Trypanosoma brucei* and *Trypanosoma cruzi*. BMC Genomics *6*, 127.

Poon, S. K., Peacock, L., Gibson, W., Gull, K., and Kelly, S. (2012). A modular and optimized single marker system for generating *Trypanosoma brucei* cell lines expressing T7 RNA polymerase and the tetracycline repressor. Open Biol *2*, 110037.

Siegel, T. N., Hekstra, D. R., Kemp, L. E., Figueiredo, L. M., Lowell, J. E., Fenyo, D., Wang, X., Dewell, S., and Cross, G. A. M. (2009). Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. Genes Dev. *23*, 1063–1076.

Skubák, P., and Pannu, N. S. (2013). Automatic protein structure solution from weak X-ray data. Nat Commun *4*, 2777.

Stankovic, A., and Jansen, L. E. T. (2017). Quantitative Microscopy Reveals Centromeric Chromatin Stability, Size, and Cell Cycle Mechanisms to Maintain Centromere Homeostasis. Prog. Mol. Subcell. Biol. *56*, 139–162.

Studier, F. W. (2005). Protein production by auto-induction in high-density shaking cultures. Protein Expression and Purification *41*, 207–234.

Suzuki, M. (1989). SPKK, a new nucleic acid-binding unit of protein found in histone. EMBO J. *8*, 797–804.

Talbert, P. B., Bayes, J. J., and Henikoff, S. (2009). Evolution of Centromeres and Kinetochores: A Two-Part Fugue. In The Kinetochore:, P. De Wulf and W. C. Earnshaw, eds. (Springer New York), pp. 1–37. Available at: http://www.springerlink.com/content/q810p641t4242254/abstract/.

Tanifuji, G., Cenci, U., Moog, D., Dean, S., Nakayama, T., David, V., Fiala, I., Curtis, B. A., Sibbald, S. J., Onodera, N. T., et al. (2017). Genome sequencing reveals metabolic and cellular interdependence in an amoeba-kinetoplastid symbiosis. Sci Rep *7*, 11688.

Terwilliger, T. C., Dimaio, F., Read, R. J., Baker, D., Bunkóczi, G., Adams, P. D., Grosse-Kunstleve, R. W., Afonine, P. V., and Echols, N. (2012). phenix.mr_rosetta: molecular replacement and model rebuilding with Phenix and Rosetta. J. Struct. Funct. Genomics *13*, 81–90.

UniProt Consortium (2019). UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. *47*, D506–D515.

Vonrhein, C., Blanc, E., Roversi, P., and Bricogne, G. (2007). Automated structure solution with autoSHARP. Methods Mol. Biol. *364*, 215–230.

Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T., and Bricogne, G. (2011). Data processing and analysis with the autoPROC toolbox. Acta Crystallogr. D Biol. Crystallogr. *67*, 293–302.

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., and Barton, G. J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics *25*, 1189–1191.

Wickstead, B., and Gull, K. (2011). The evolution of the cytoskeleton. J. Cell Biol. *194*, 513–525.

Winter, G. (2010). xia2: an expert system for macromolecular crystallography data reduction. J Appl Cryst *43*, 186–190.

Winter, G., Waterman, D. G., Parkhurst, J. M., Brewster, A. S., Gildea, R. J., Gerstel, M., Fuentes-Montero, L., Vollmar, M., Michels-Clark, T., Young, I. D., et al. (2018). DIALS: implementation and evaluation of a new integration package. Acta Cryst D *74*, 85–97.

Wolfe, S. A., Nekludova, L., and Pabo, C. O. (2000). DNA recognition by Cys2His2 zinc finger proteins. Annu Rev Biophys Biomol Struct *29*, 183–212.

Yu, B., Martins, I. R. S., Li, P., Amarasinghe, G. K., Umetani, J., Fernandez-Zapico, M. E., Billadeau, D. D., Machius, M., Tomchick, D. R., and Rosen, M. K. (2010). Structural and energetic mechanisms of cooperative autoinhibition and activation of Vav1. Cell *140*, 246–256.

Table 1

Data collection	*Bs*KKT2 central domain
Beamline	Diamond I04
Wavelength (Å)	1.28297
Space group (Z)	I222 (8)
Unit cell (cell edges in Å, cell angles in degrees)	38.96, 53.51, 83.29 90, 90, 90
Resolution range (Å)	45.02 – 1.8 (1.83 – 1.8)
Total No of reflections	37266 (925)
No. of unique reflections	8158 (334)
Completeness (%)	96.96 (77.86)
R_{merge}	0.10 (0.52)
R_{pim}	0.05 (0.33)
$CC_{1/2}$	0.99 (0.77)
$<I/\sigma(I)>$	8.65 (1.08)
Multiplicity	4.57 (2.77)
Anomalous completeness (%)	89.83 (37.37)
Anomalous multiplicity	2.54 (1.93)
Overall *B* factor from Wilson plot ($Å^2$)	22.4
Refinement	
Resolution (Å)	35.3 – 1.8 (2.0 – 1.8)
No. of reflections working set	7748 (390)
No. of reflections test set	403 (18)
Final R_{work} (%)	19.5 (24.3)
Final R_{free} (%)	22.5 (23.4)
No. of protein atoms	770
No. of Zn atoms	3
No. of water atoms	94
No. of sulfate ions	2
Average B factor ($Å^2$) protein atoms	30.45
Average B factor ($Å^2$) Zn atoms	25.31
Average B factor ($Å^2$) water atoms	44.43
Rmsd bond lengths (Å)	0.008
Rmsd bond angles (°)	0.96
Ramachandran plot	
Most favoured (%)	97.89
Allowed (%)	2.11
Disallowed	0

In parentheses the values relative to the highest resolution shell.

Table 2

Data collection	*Pk*KKT2a central domain	
Beamline	Diamond I24	Diamond I03
Wavelength (Å)	0.96861	1.28272
Space group (Z)	$P6_4$ (6)	$P6_4$ (6)
Unit cell (cell edges in Å, cell angles in degrees)	113.84, 113.84, 46.01, 90, 90, 120	114.33, 114.33, 46.32 90, 90, 120
Resolution range (Å)	56.92 – 2.87 (2.92 - 2.87)	99.01 – 3.80 (3.86 - 3.80)
Total No of reflections	153076 (6844)	68026 (3882)
No. of unique reflections	7986 (357)	3535 (189)
Completeness (%)	99.73 (91.77)	100.00 (100.00)
R_{merge}	0.10 (1.04)	0.31 (6.13)
R_{pim}	0.024 (0.240)	0.073 (1.38)
$CC_{1/2}$	0.99 (0.45)	0.99 (0.4)
$<I/\sigma(I)>$	16.45 (2.14)	12 (1.7)
Multiplicity	19.17 (19.17)	19.2 (20.5)
Anomalous completeness (%)	99.6 (91.2)	100 (100)
Anomalous multiplicity	10 (9.9)	10.2 (10.6)
Overall *B* factor from Wilson plot ($Å^2$)	105.95	136.7
Refinement		
Resolution (Å)	49 – 2.87 (3 – 2.87)	
No. of reflections working set	7408 (417)	
No. of reflections test set	377 (16)	
Final R_{work} (%)	25.2 (36.6)	
Final R_{free} (%)	27.5 (51.2)	
No. of protein atoms	832	
No. of Zn atoms	2	
Average B factor ($Å^2$) protein atoms	113.4	
Average B factor ($Å^2$) Zn atoms	100.6	
Rmsd bond lengths (Å)	0.009	
Rmsd bond angles (°)	1.11	
Ramachandran plot		
Most favoured (%)	90.9	
Allowed (%)	9.1	
Disallowed	0	

In parentheses the values relative to the highest resolution shell.

Figures

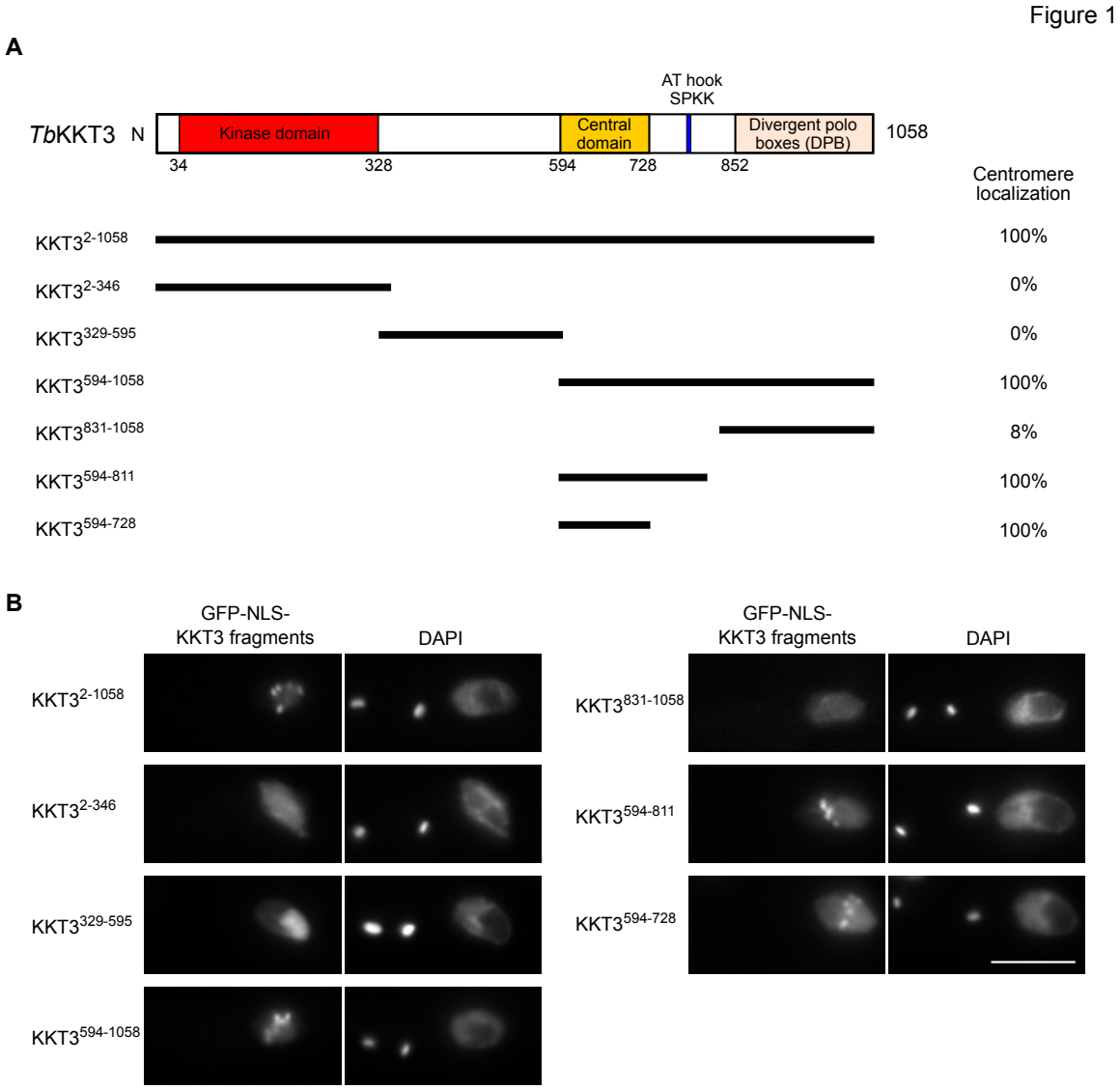
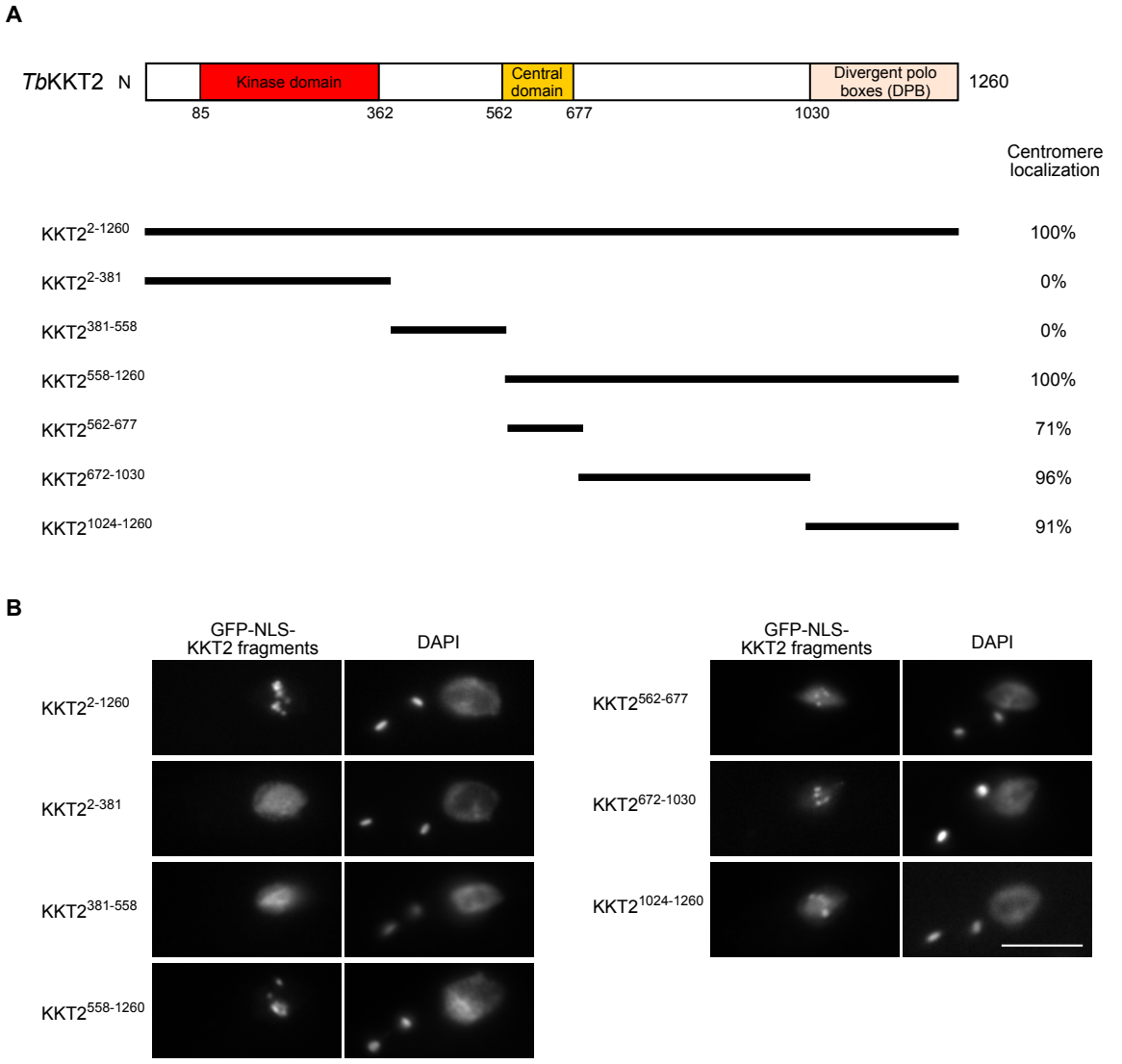


Figure 1. KKT3 central domain is able to localize at centromeres in *T. brucei*

(A) Schematic of the *T. brucei* KKT3 protein. Percentages of GFP-positive 2K1N cells that have kinetochore-like dots were quantified at 1 day post-induction (n > 24, each).

(B) Ectopically expressed *Tb*KKT3 fragments that contain the central domain form kinetochore-like dots. Inducible GFP-NLS fusion proteins were expressed with 10 ng/mL doxycycline. Cell lines, BAP291, BAP292, BAP379, BAP296, BAP378, BAP377, BAP418. Scale bar, 5 µm.

Figure 2



C

Proteins co-purified with KKT2$^{1024-1260}$	WT Mascot score	WT Coverage	WT Number of peptides	W1048A Number of peptides
KKT2	2341	22.1%	106	116
KKT1	458	9.8%	22	0
KKT6	105	23.4%	3	0
KKT7	96	7.9%	4	0
KKT8	80	17.5%	7	0

D

Proteins co-purified with KKT2$^{672-1030}$	Mascot score	Coverage	Number of peptides
KKT2	6263	23.7%	343
KKIP1	2574	46.4%	126
KKT3	368	14.3%	18
KKT1	199	6.1%	10
KKT7	144	10.4%	8
KKT8	75	15.8%	4
KKT10	71	3.9%	2

Figure 2. KKT2 has multiple domains that can promote centromere localization in *T. brucei*

(A) Schematic of the *T. brucei* KKT2 protein. Percentages of GFP-positive 2K1N cells that have kinetochore-like dots were quantified at 1 day post-induction (n > 22, each).

(B) Ectopically expressed *Tb*KKT2 fragments that contain either the central domain (562–677), 672–1030, or the divergent polo boxes (1024–1260) form kinetochore-like dots. Inducible GFP-NLS

fusion proteins were expressed with 10 ng/mL doxycycline. Cell lines, BAP327, BAP328, BAP381, BAP331, BAP457, BAP519, BAP517. Scale bar, 5 µm.

(C) *Tb*KKT2 DPB WT, not W1048A, co-purifies with several kinetochore proteins. Inducible GFP-NLS fusion proteins were expressed with 10 ng/mL doxycycline, and immunoprecipitation was performed using anti-GFP antibodies. Cell lines, BAP517, BAP535.

(D) *Tb*KKT2$^{672–1030}$ co-purifies with KKIP1 and several kinetochore proteins. Cell line, BAP519.

Figure 3

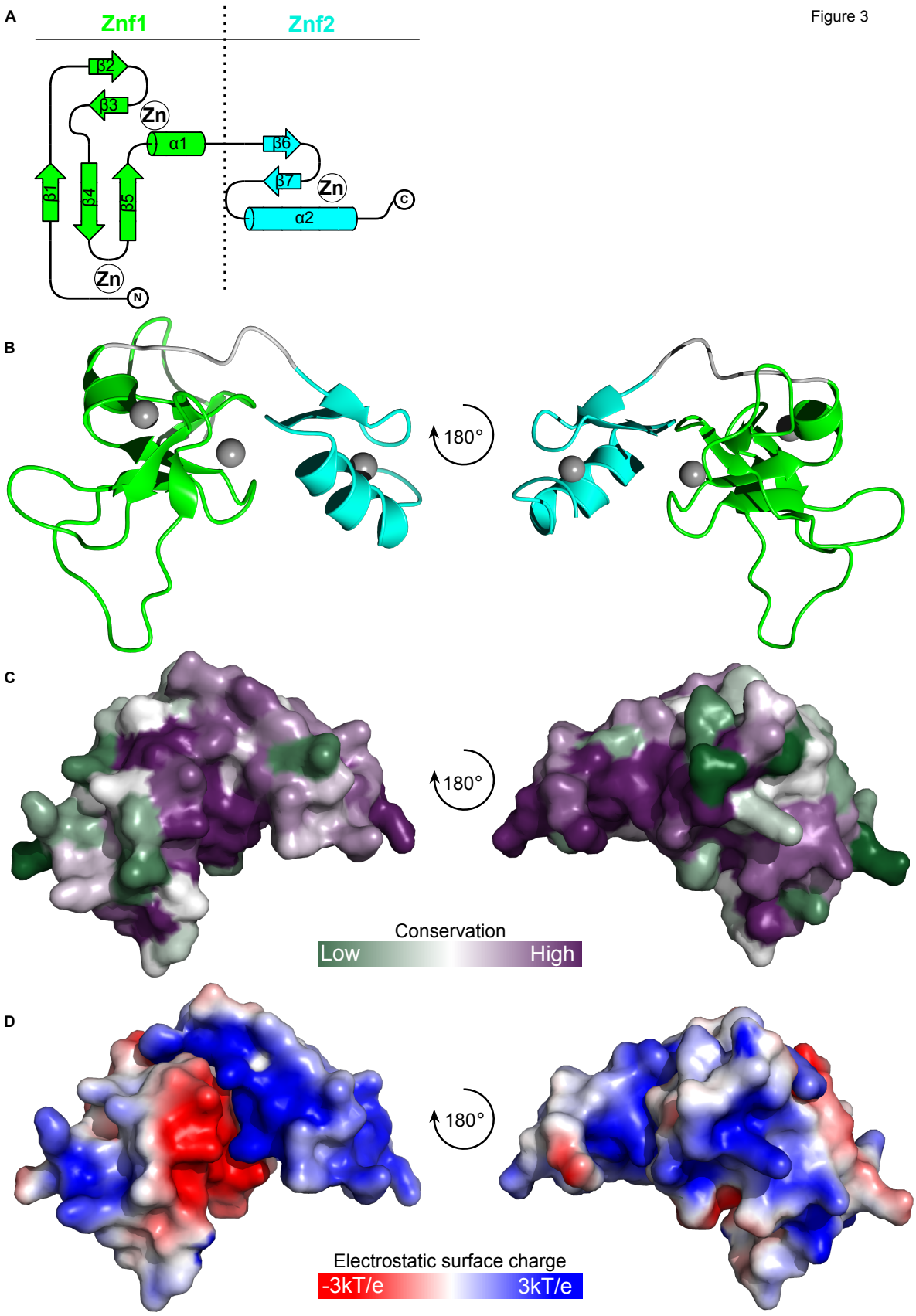

A
Znf1
Znf2
β2
β3
Zn
α1
β6
β7
Zn
α2
C
β1
β4
β5
Zn
N
B
180°
C
180°
Conservation
Low
High
D
180°
Electrostatic surface charge
-3kT/e
3kT/e

Figure 3. Crystal structure of *Bodo saltans* KKT2 central domain reveals the presence of two zinc fingers Znf1 and Znf2

(A) Topology diagram of the *Bs*KKT2 central domain showing Znf1 in green and Znf2 in cyan.

(B) Cartoon representation of the *Bs*KKT2 central domain in two orientations. Zinc ions are shown in grey spheres. The structure is colored as in (A).

(C) Surface representation of the *Bs*KKT2 central domain colored according to sequence conservation using the ConSurf server (Landau et al., 2005; Ashkenazy et al., 2016). Structure orientation as in (B).

(D) Electrostatic surface potential of the *Bs*KKT2 central domain generated by APBS (Jurrus et al., 2018). Structure orientation as in (B).

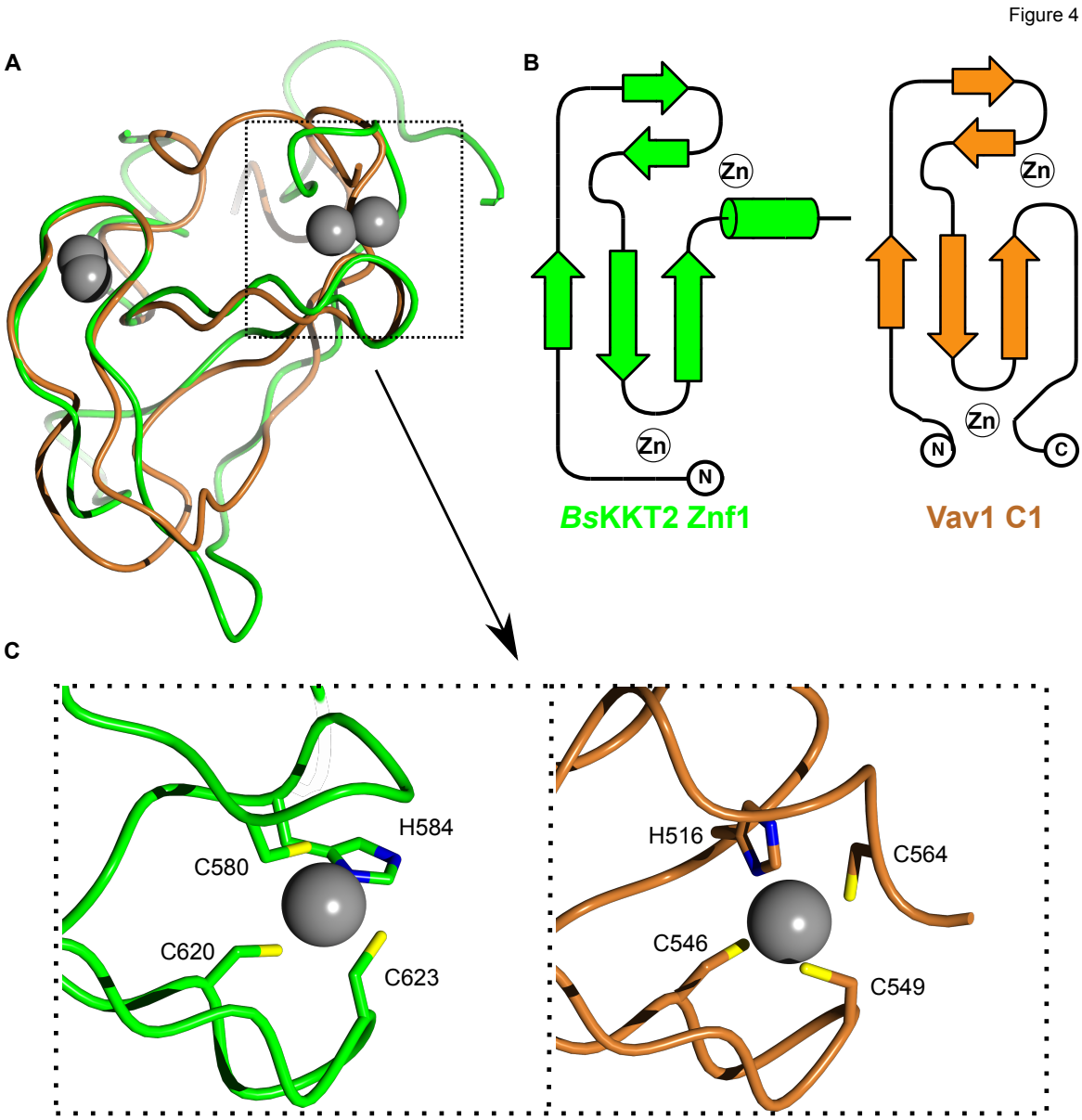


Figure 4. *Bodo saltans* KKT2 Znf1 is a unique domain

(A) Structure superposition of *Bs*KKT2 Znf1 in green and its structural homolog Vav1 C1 domain in brown (PDB: 3KY9 (Yu et al., 2010)). Zinc ions are shown in grey spheres.

(B) Topology diagram of *Bs*KKT2 Znf1 and Vav1 C1 domain.

(C) Close-up view showing a key difference in zinc coordination between *Bs*KKT2 Znf1 and Vav1 C1 domain.

Figure 5

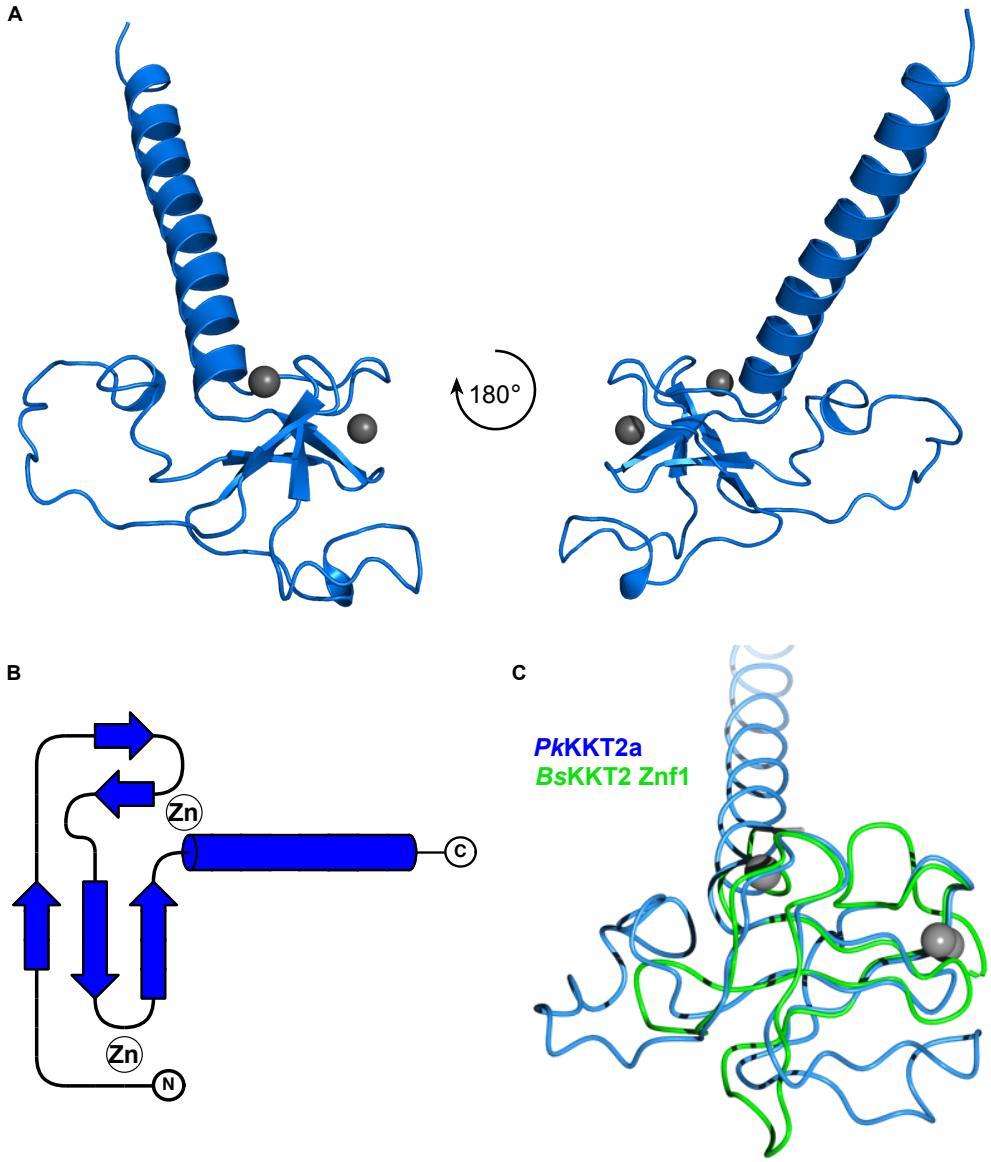


Figure 5. Crystal structure of *Perkinsela* KKT2a central domain highlights the conservation of Znf1

(A) Cartoon representation of the *Pk*KKT2a central domain in two orientations. Zinc ions are shown in grey spheres.

(B) Topology diagram of the *Pk*KKT2a central domain structure.

(C) Structure superposition of the *Pk*KKT2a central domain and *Bs*KKT2 Znf1, showing that the core of the zinc finger structure is conserved. Variations between the two structures are due to sequence insertions within Znf1 and the absence of the Znf2 domain at the C-terminus in the *Pk*KKT2a central domain.

Figure 6

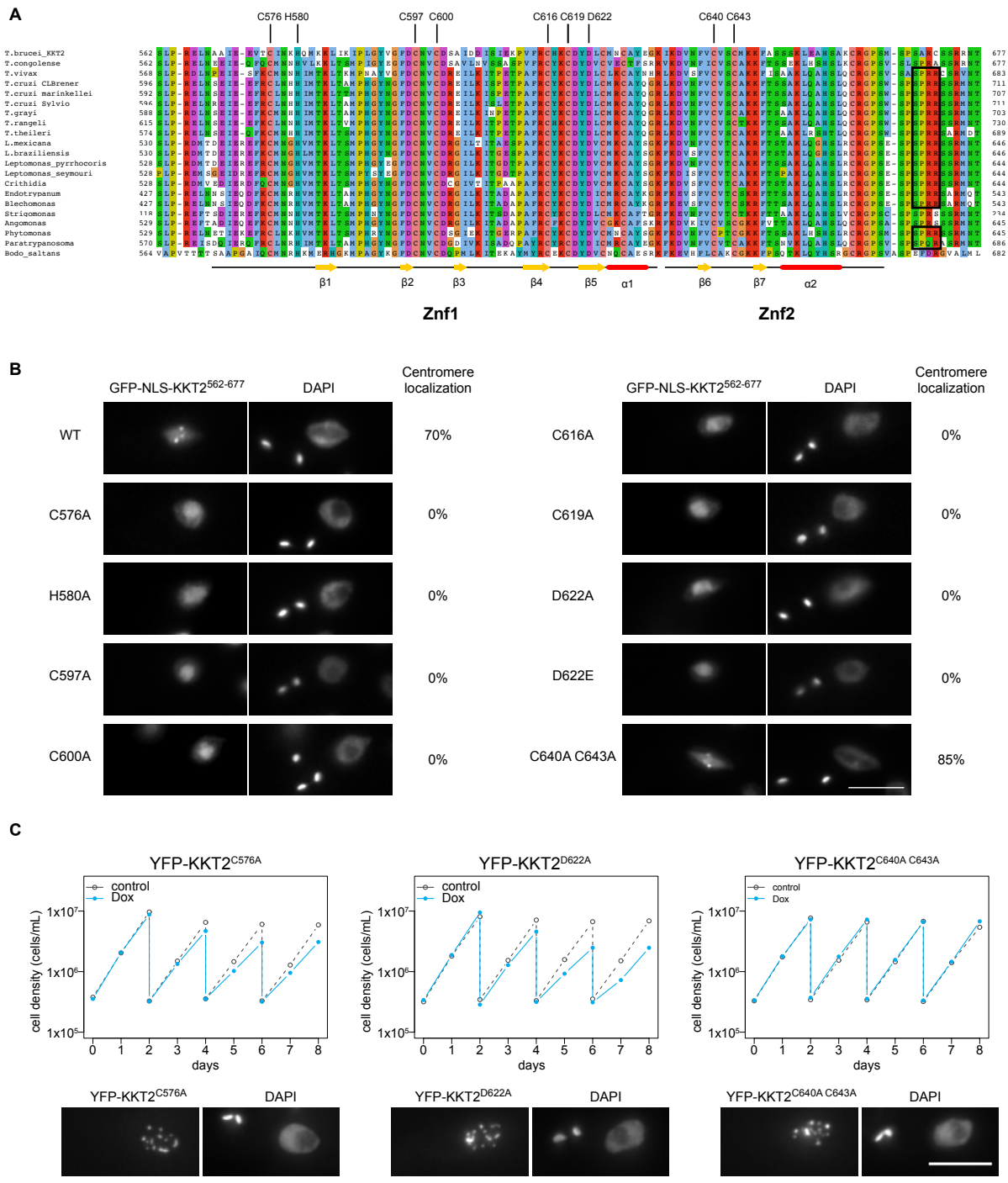


Figure 6. KKT2 Znf1 is critical for the centromere localization *in T. brucei*

(A) Multiple sequence alignment of KKT2. Residues mutated in (B) as well as secondary structures of *Bs*KKT2 are shown. Putative SPKK motifs are highlighted in boxes.

(B) Kinetochore localization of *Tb*KKT2$^{562–677}$ depends on Znf1, not Znf2. Percentages of GFP-positive 2K1N cells that have kinetochore-like dots were quantified at 1 day post-induction (n = 40,

each). Inducible GFP-NLS fusion proteins were expressed with 10 ng/mL doxycycline. Cell lines, BAP457, BAP1700, BAP1702, BAP1710, BAP1712, BAP1715, BAP1717, BAP1649, BAP1719, BAP1837.

(C) *Tb*KKT2 C576A and D622A mutants localize at kinetochores but fail to support normal cell growth. One allele of *Tb*KKT2 was mutated and tagged with an N-terminal YFP tag, and the other allele was depleted using RNAi-mediated knockdown by targeting the 5'UTR of the *Tb*KKT2 transcript. Top: cells were diluted every two days and cell growth was monitored for 8 days upon induction of RNAi. Controls are uninduced cell cultures. Similar results were obtained for at least three clones of *Tb*KKT2 mutants. Bottom: Example of cells expressing the *Tb*KKT2 mutants prior to RNAi induction, showing that they localize at kinetochores (n > 150, each). Maximum intensity projection are shown. RNAi was induced with 1 µg/mL doxycycline. Cell lines, BAP1789, BAP1779, BAP1786. Scale bars, 5 µm.

Figure 7

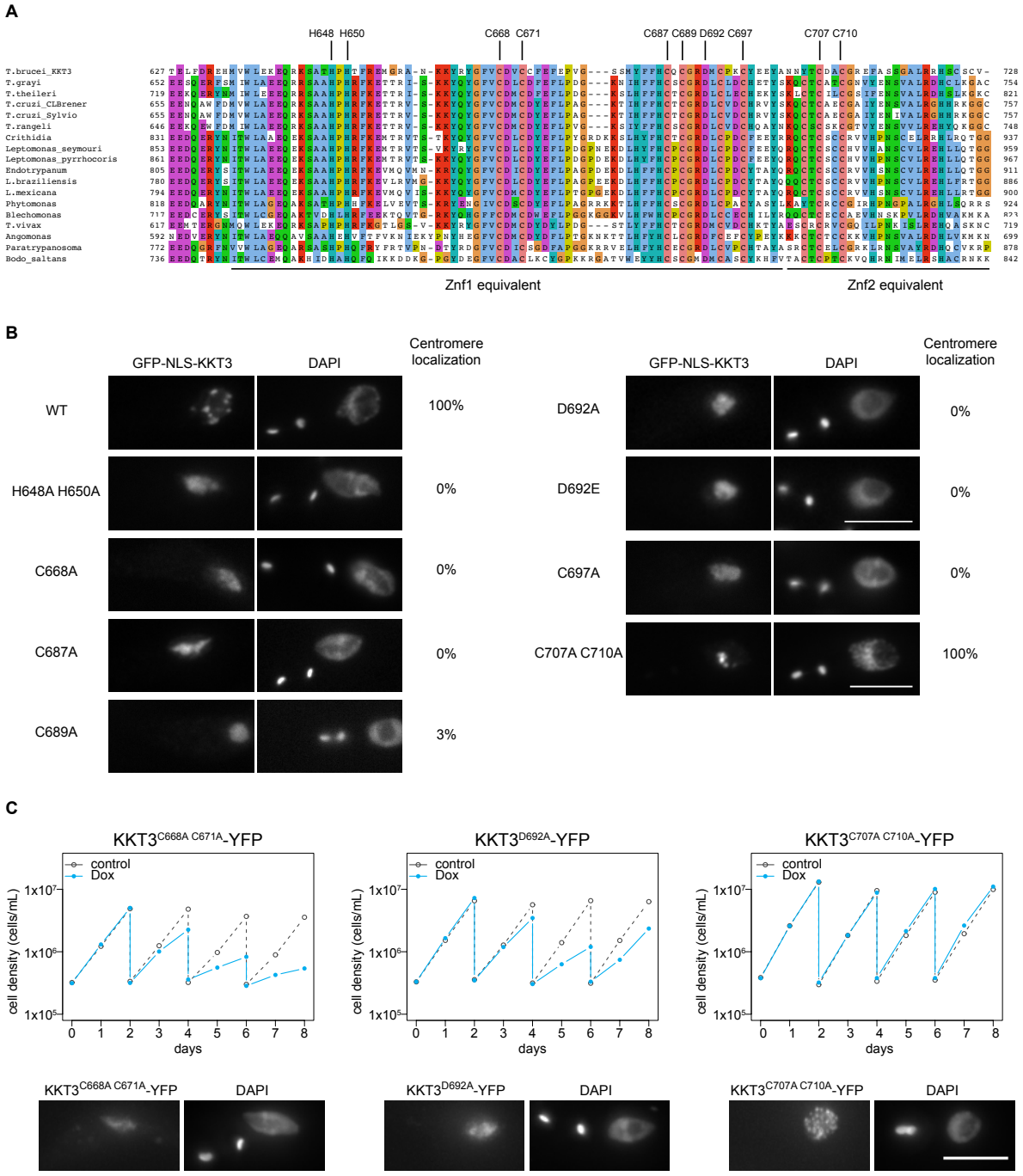


Figure 7. Kinetochore localization of KKT3 depends on the central domain in *T. brucei*

(A) Multiple sequence alignment of KKT3. Residues mutated in (B) or (C) are shown.

(B) Percentage of GFP-positive cells that have kinetochore-like dots were quantified at 1 day post-induction (n > 22, each). Inducible GFP-NLS fusion proteins were expressed with 10 ng/mL

doxycycline. Cell lines, BAP291, BAP359, BAP360, BAP446, BAP447, BAP1721, BAP1722, BAP362, BAP341.

(C) *Tb*KKT3 C668A/C671A and D692A mutants do not localize at kinetochores and fail to support normal cell growth, while *Tb*KKT3 C707A/C710A mutant is functional. One allele of *Tb*KKT3 was mutated and tagged with a C-terminal YFP tag, and the other allele was depleted using RNAi-mediated knockdown by targeting the 3'UTR of the *Tb*KKT3 transcript. Top: cells were diluted every two days and cell growth was monitored for 8 days upon induction of RNAi. Similar results were obtained for at least three clones of *Tb*KKT3 mutants. Controls are uninduced cell cultures. Bottom: Example of cells expressing the *Tb*KKT3 mutants prior to RNAi induction, showing that *Tb*KKT3$^{C668A\ C671A}$ and *Tb*KKT3^{D692A} do not localize at kinetochores while *Tb*KKT3$^{C707A\ C710A}$ localizes normally (n > 90, each). Maximum intensity projection are shown. RNAi was induced with 1 µg/mL doxycycline. Cell lines, BAP1791, BAP1793, BAP1783. Scale bars, 5 µm.

Supplemental material

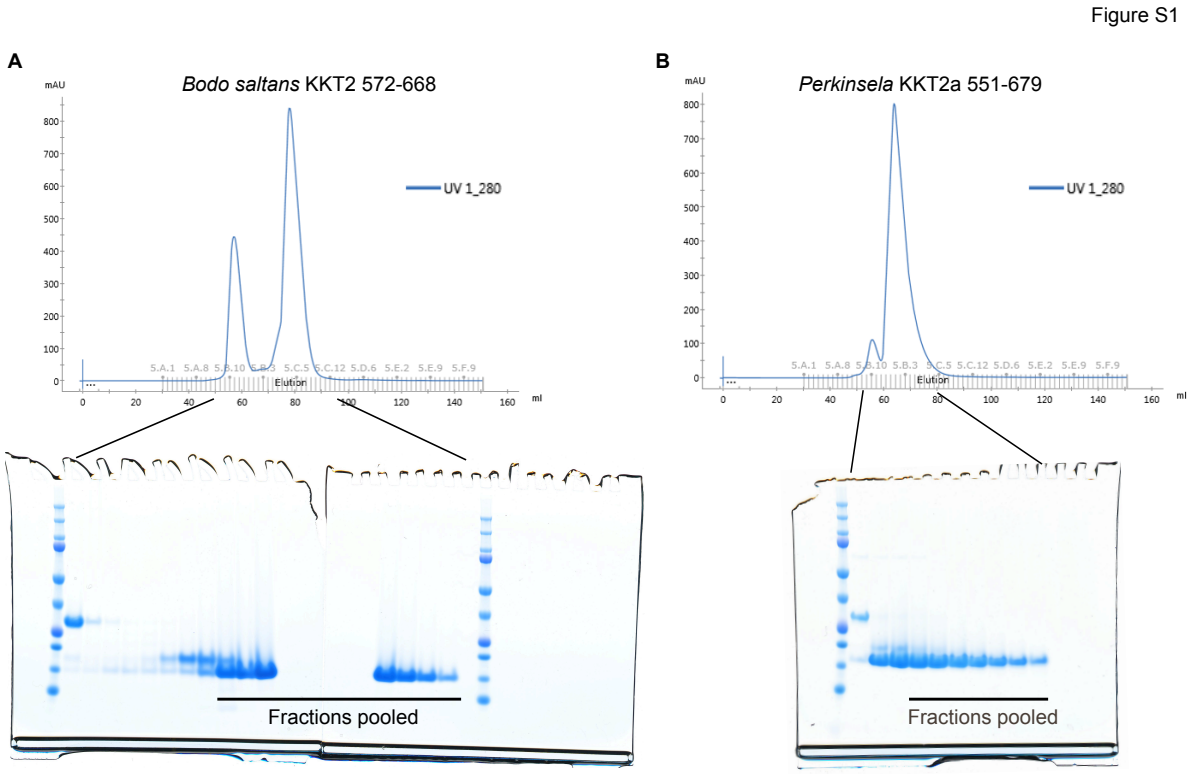


Figure S1. Purification of *Bodo saltans* KKT2 and *Perkinsela* KKT2a

Size exclusion chromatography of (A) *Bodo saltans* KKT2 central domain and (B) *Perkinsela* KKT2a central domain with respective SDS-PAGE gels showing pooled fractions.

Figure S2

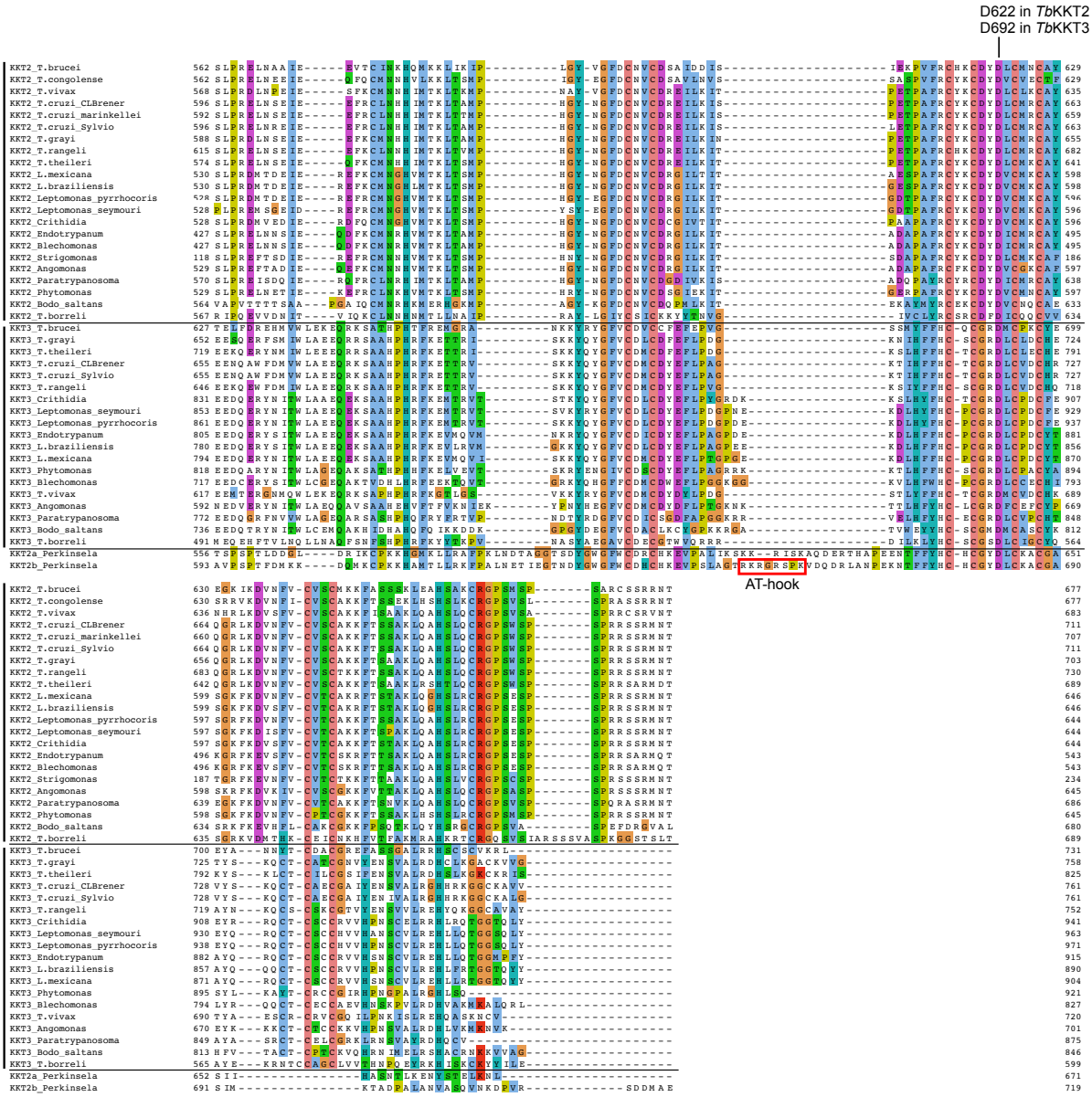


Figure S2. Multiple sequence alignment of KKT2 and KKT3 from various kinetoplastids, as well as *Perkinsela* KKT2a and KKT2b

Position of the conserved aspartic acid is highlighted. A putative AT-hook motif in *Perkinsela* KKT2b is highlighted in a red box.

Figure S3

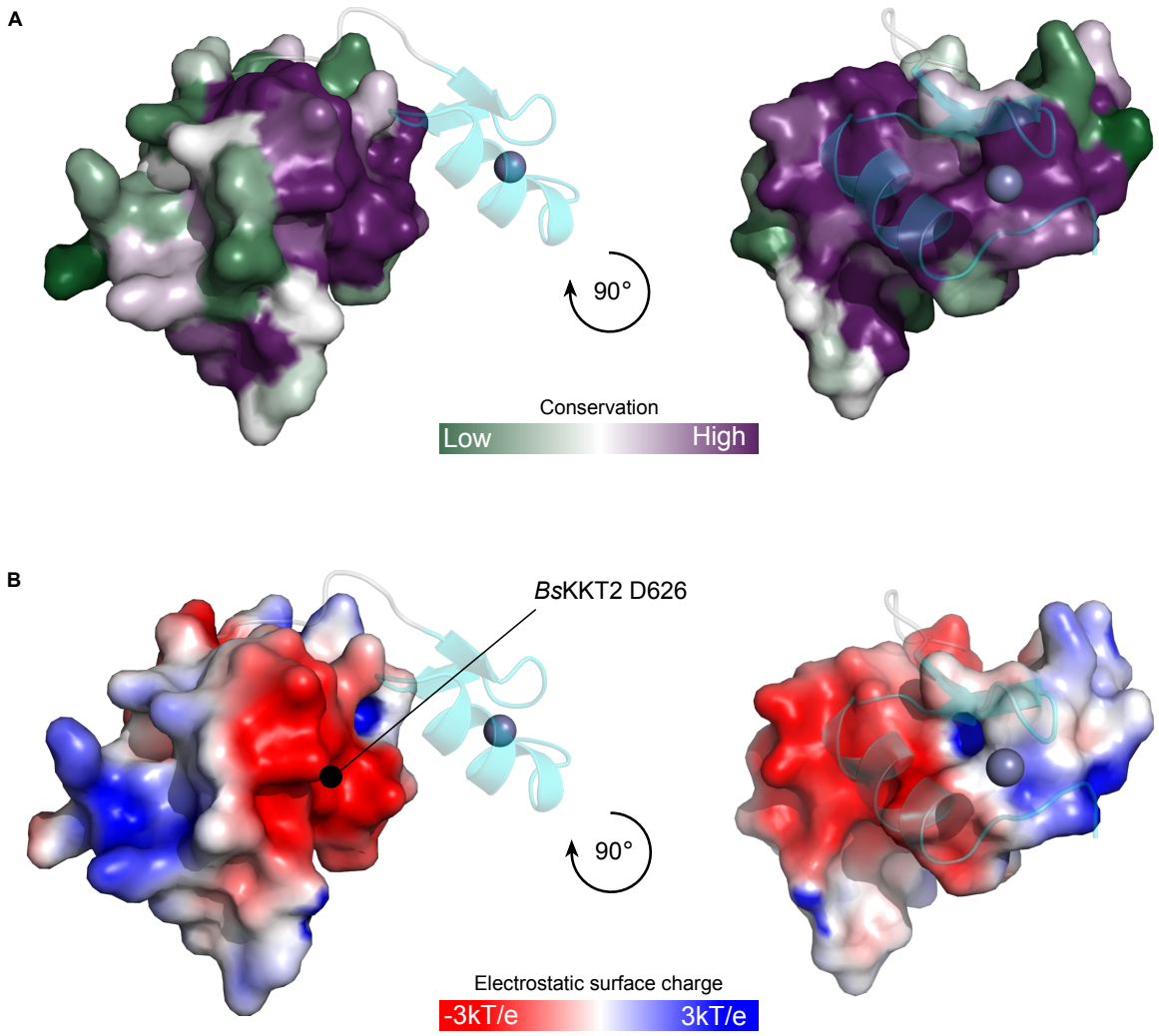


Figure S3. *Bodo saltans* KKT2 Znf1 surface sequence conservation and electrostatic potential reveal the presence of a conserved acidic patch centering on residue D626

(A) Surface sequence conservation of *Bs*KKT2 Znf1 using the ConSurf server (Landau et al., 2005; Ashkenazy et al., 2016). Znf2 structure is shown as cartoon representation.

(B) Electrostatic surface potential of *Bs*KKT2 Znf1 generated by APBS (Jurrus et al., 2018) reveals the presence of a conserved acidic surface in Znf1. The location of the residue D626 is marked by a black circle.

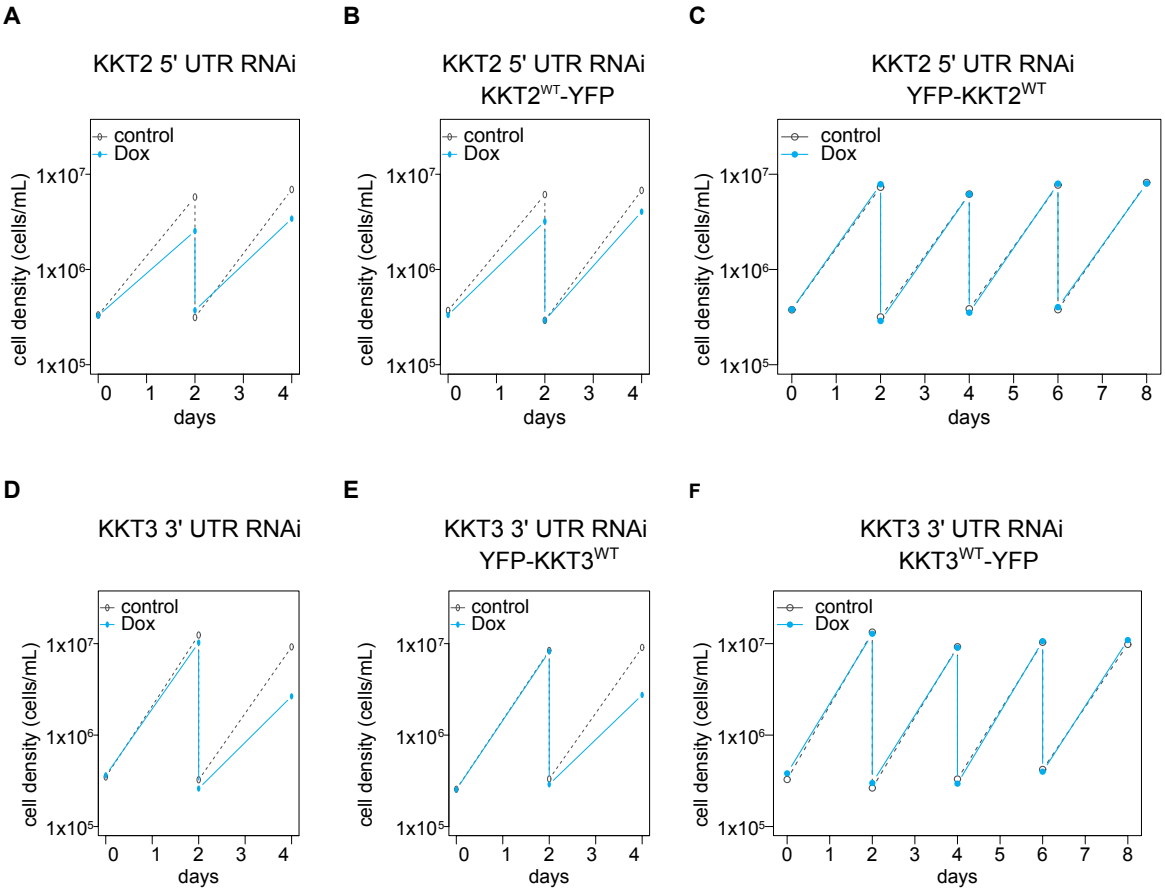


Figure S4. RNAi of KKT2 and KKT3 causes growth defects in *T. brucei*

Growth curve of (A) KKT2 5'UTR RNAi, (B) KKT2-YFP, KKT2 5'UTR RNAi, (C) YFP-KKT2, KKT2 5'UTR RNAi, (D) KKT3 3'UTR RNAi, (E) YFP-KKT3, KKT3 3'UTR RNAi, and (F) KKT3-YFP, KKT3 3'UTR RNAi. 1 µg/mL doxycycline was added to induce RNAi. Controls are uninduced cell cultures. Similar results were obtained from at least two independent experiments. Cell lines, BAP1554, BAP1752, BAP1681, BAP1555, BAP1659, BAP1682.